Advances in Artificial Intelligence Research (AAIR)
ISSN 2757-7422
www.dergipark.com/aair/

Vol. 3 (No. 2), pp. 85-95, 2023
doi: 10.54569/aair.1254810
Published online: Oct 23, 2023

A Machine Learning Approach for Simultaneous Classification of Material Types and Cracks

Ömer Mintemur [1, *]

[1] Ankara Yıldırım Beyazıt University School of Engineering and Natural Sciences

Abstract

Exterior structures are susceptible to deformation, which can manifest as cracks on the surface. Deformations that occur on surfaces subjected to daily human use can exacerbate rapidly, potentially leading to irreversible structural damage. They have a potential to result in fatalities. Thus, continuous inspection of these deformations is of invaluable importance. In addition, the identification of the materials comprising the structures is essential to facilitate the implementation of appropriate precautionary measures. However, the inspections are hard to maintain with a solely human workforce. More advanced actions can be taken thanks to the developments in technology. Machine Learning methods could be used in this area where human workforce is ineffective. In this regard, an end-to-end Machine Learning approach was proposed in this study. The power of classical feature extraction methods and Artificial Neural Networks were combined to detect cracks and material of the surface simultaneously. The 2D Discrete Wavelet Transform and statistical properties gained from Gray Level Co-Occurrence Matrix were utilized in the feature extraction mechanism, and an ANN structure was designed. The findings of the study indicate that the proposed mechanism achieved an acceptable level of accuracy for recognizing the structural deformations, despite the challenges posed by the complexity of the problem.

Keywords: *Classification; material recognition; machine learning; crack detection.*

1. Introduction

Structures are susceptible to deformation (*cracks*) over an extended period of time and with frequent use [1,2]. The most common symptom of deformation is the appearance of cracks on the surface of the structure. A crack is a type of discontinuity that can form on the surface of a material, characterized by a break or separation along a portion of the surface, disrupting its uniformity and integrity [3]. The occurrence of surface cracks can have a considerable impact on the strength, stability, and aesthetic appearance of the material. In certain instances, surface cracks may also be indicative of deeper structural issues, requiring further investigation [4].

The structural deterioration tends to occur with greater frequency on exteriors such as walls, roads, and pavements. An example of the progression of surface degradation can be seen in the transformation of cracks on roads into holes over time [5]. The presence of holes on pavements presents a significant threat to public safety [6]. As such, it is essential to continuously monitor and promptly address these structural deficiencies to ensure the longevity and safety of the infrastructure.

However, conducting ongoing monitoring is not practical due to the constraints of the available workforce [7]. But advancements in technology allow for the delegation of ongoing monitoring tasks to computers. The utilization of computer vision techniques has become increasingly prevalent in the pursuit of continuous surface defect detection.

The established methodology for this task entails the extraction of salient features from an image, followed by an algorithm that categorizes the image as either exhibiting defects or being free from defects. The generation of features constitutes a component of the broader domain of image processing, encompassing various techniques including edge detection, extraction of color information, thresholding, and calculation of statistical characteristics such as energy and contrast etc. [8,9]. On the other hand, classification algorithms are situated within the domain of Artificial Intelligence (AI) and are commonly known as Machine Learning (ML) algorithms. The extracted features are input into an ML algorithm, which performs the classification.

There exists a substantial body of research that employs these conventional methodologies. An end-to-end methodology for identifying cracks in asphalt surfaces can be found in [10]. The methodology was divided into several sections. Prior to the feature extraction, thresholding and noise elimination were employed to enhance the quality of the acquired image. In the feature extraction phase, the Hough Transform (HT) [11] was utilized to extract features from the image. Finally, Bagging, which is an ensemble ML method [12] and Support Vector Machines (SVM) [13] were used to classify the cracks.

A study utilized Local Binary Pattern (LBP) [14] and Principal Component Analysis (PCA) [15] to identify

*Corresponding author
E-mail address: omintemur@ybu.edu.tr
Received: 22/Feb/2023; Received in revised form: 09/Aug/2023; Accepted: 01/Oct/2023.

cracks on the pavements can be found in [16]. The extracted features were classified using SVM, resulting in a high rate of classification accuracy, as indicated by the authors.

Statistical information extracted from a given image can serve as a significant feature for the purpose of classification, and such a study used these numerical values to detect the defect in pavements [17]. The authors used statistical properties that are calculated by constructing Gray-Level Co-Occurrence Matrix (GLCM) [18]. The implementation of this approach resulted in a substantial classification rate, estimated to be approximately 88%.

In addition to the conventional image processing (CIP) techniques, there has been a recent trend in the field towards the utilization of Convolutional Neural Networks (CNNs) as a means for image classification. CNNs can be classified as a sub class of Artificial Neural Networks (ANNs) and they offer a unique advantage as an end-to-end system, demonstrating exceptional performance in both feature extraction and classification due to its utilization of the backpropagation algorithm [19,20]. One of the key advantages of using CNNs in crack classification is the elimination of the manual feature extraction process. And these powerful networks have been widely utilized in the literature of crack detection [3, 21-24].

However, both CNNs and CIP techniques in this area have their own limitations that can be extended. The process of feature extraction in CNNs is automated and has the possibility to identify redundant features that may impede the efficiency of the training procedure. Also, it is noteworthy that there is currently no universally accepted method for the construction of these networks.

On the other hand, CIP techniques that have been proposed in this field are limited by their singular approach. Nevertheless, there exist potent image processing methods that can be combined to achieve more robust feature from the image data. Furthermore, classical ML algorithms are also limited by their finite number of parameters and are therefore not well-suited for modifications.

Therefore, a research inquiry emerges as to whether the capabilities of two sub-fields (CIP and ANN) of AI can be combined to produce competitive results in the area of crack detection.

The current research focus in this domain primarily centers around the detection of cracks on surfaces. However, the exclusive monitoring of cracks may not be beneficial in overall. Classifying the type of surfaces is also a crucial factor in ensuring appropriate maintenance and emergency response measures. This raises a second research question as to whether it is feasible to detect both cracks and the surface material with high accuracy.

This paper presents an end-to-end ML approach for detecting cracks and classifying the type of surface in question at the same time, in accordance with the mentioned research questions. The images of cracks on a surface tend to exhibit a distinctive linear orientation, which may be characterized as horizontal, vertical, or diagonal. These orientations were enhanced by using Discrete Wavelet Transform (DWT). GLCM features were utilized to extract statistical features from the enhanced images [17,25]. Finally, the obtained features were fed into an ANN structure to detect both cracks and surface material. The experimental evaluations were conducted using an image dataset produced by Utah State University, as it offers a diverse representation of various materials along with both cracked and non-cracked images.

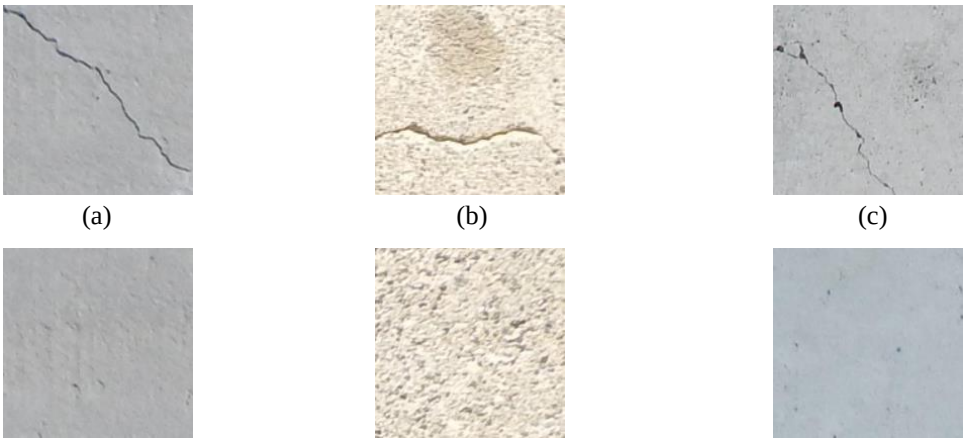
The rest of the paper is organized as follows: The details regarding the dataset utilized and the methodology employed comprehensively described in Section 2. Section 3 outlines the experimental setup, configurations, and metrics employed to evaluate the efficacy of the proposed method. The results of the experiments, along with their interpretation, are presented in Section 4. Finally, the concluding remarks and suggestions for future work are presented in the last section of the paper.

2. Materials and Methods

This section briefly gives information to the reader about the dataset, the methods utilized, and the experimental setup in this paper.

2.1. Dataset

The present study utilized the SDNET2018 dataset to evaluate the efficacy of the proposed methodology. The dataset comprises three distinct surface types, namely, decks, walls, and pavements, each of which is categorized into two classes based on the presence or absence of cracks. The dataset consists of a total of 56,000 images depicting cracked and non-cracked decks, walls, and pavements, each with a resolution of 256x256 pixels. Total number of cracked images is 8484 and non-cracked images is 47607 [26,27]. Each image in the dataset is presented in RGB color format. Straightforward exemplary images from each class are given in **Figure 1**.

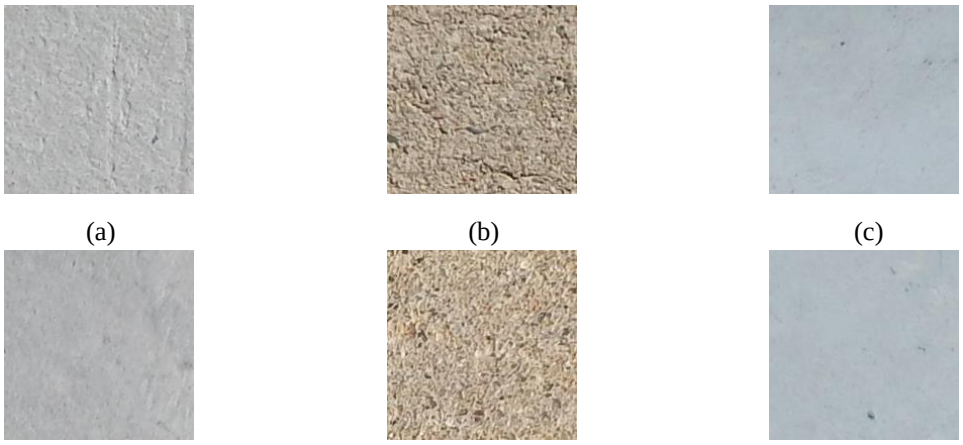


(a) (b) (c)

(d) (e) (f)

Figure 1. *Example Images From the SDNET2018. (a) Decks – Cracked (b) Pavements – Cracked (c) Walls – Cracked (d) Decks – Non Cracked (e) Pavements – Non Cracked (f) Walls – Non Cracked (Figure is in color in online version of paper)*

Upon examination of **Figure 1**, it becomes apparent that cracked and non-cracked images exhibit discernible patterns that differentiate them from one another. Nevertheless, the dataset includes images that exhibit considerable complexity and challenge even to human observers. Such examples are illustrated in **Figure 2**.



(a) (b) (c)

(d) (e) (f)

Figure 2. *Images with Considerable Complexity from the SDNET2018. (a) Decks – Cracked (b) Pavements – Cracked (c) Walls – Cracked (d) Decks – Non-Cracked (e) Pavements – Non-Cracked (f) Walls – Non-Cracked (Figure is in color in online version of paper)*

Cracked and non-cracked images exhibit differences in their linear or non-linear orientation perspectives. Additionally, these orientations reveal themselves at different angles. As can be seen from the example images, these orientations exhibit strong pattern changes (*edges*) in the images. These pattern changes could be enhanced, in other words, a mechanism could be employed to present strong pattern changes in these images. Thus, DWT was employed to enrich those pattern changes. The next subsection describes the DWT briefly and presents an example image which DWT applied.

2.2. 2D - Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is a decomposition technique that enables the analysis of 1D and 2D signals in different frequency components. 2D-DWT allows for inspection of any image in multi-resolution. The 2D-DWT provides low and high frequency information about an image at different decomposition levels. In general, 2D-DWT divides images into four different frequency bands and each of them reveals different information about a given image:

1. LL – Low Frequency component of the image
2. LH – Horizontal Edges Enhanced
3. HL – Vertical Edges Enhanced

4. HH – Diagonal Edges Enhanced

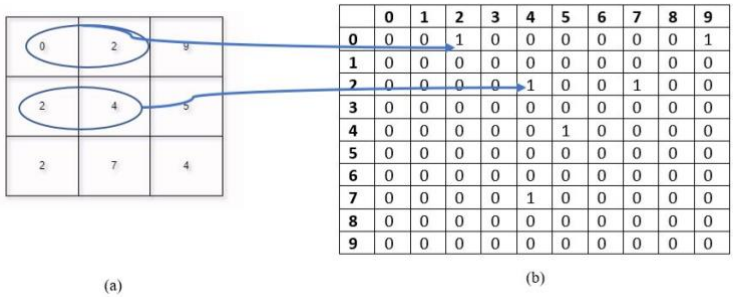
While 1-Level of decomposition provides information about 3 different orientations, one advantage of 2D DWT is that it enables *n*-Level decomposition by decomposing the LL part of the image at each level, thus providing more substantial information about the image. The 2D-DWT algorithm initiates by selecting a wavelet basis function for the image at hand. The selected wavelet function is then used to decompose the image into its component frequency sub-bands. Given the long-standing history of 2D-DWT, numerous wavelet functions have been proposed in the literature, each of which enhances the LH, HL, and HH sub-bands differently [28]. An example 4-Level decomposition of a crack deck surface image is given in **Figure 3**.

Figure 3. *4-Level Decomposition of Cracked Deck Image (Wavelet Function is Bior1.5)*

As illustrated in **Figure 3**, an increase in the level of decomposition leads to a more evident exposition of crack characteristics. Each decomposition in 2D-DWT can be used as raw features. However, the feature vector size may become excessively large for a given image, as the feature vector expands with each level of decomposition, in proportion to the image's shape. Another drawback of using 2D-DWT's results directly is the presence of unnecessary pixel information, as not every pixel is an important feature. Therefore, more meaningful features that summarize the image information into a single number could be effective for both interpretability and computational burden. For this reason, GLCM was employed in this study. The effectiveness of GLCM has been demonstrated in the literature [17, 25]. The next subsection gives brief information about GLCM to the reader.

2.3. Gray Level Co-Occurrence Matrix (GLCM)

GLCM is a statistical method that calculates correlation between pixels in a grayscale image. It evaluates the correlation between the gray level values of two pixels at a specific distance and angle in an image. It was proposed by Haralick [18]. The GLCM performs its operation using three parameters: a grayscale image, a distance parameter that determines how many pixels will contribute to correlation, and an angle parameter for which the correlation will be sought. It is capable of capturing texture changes effectively and presents this information in a number of properties: Dissimilarity, Correlation, Homogeneity, Contrast, ASM and Energy. General working mechanism and formula of each property is given in **Figure 4** and **Table 1** respectively.



	0	1	2	3	4	5	6	7	8	9
0	0	0	1	0	0	0	0	0	0	1
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	0	0	1	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	1	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	1	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0

Figure 4.*Working Mechanism of GLCM. (a) Gray values of the image. (b) Constructed GLCM of the image.*

Table 1. *GLCM Properties and Its' Respective Formulas.*

Properties	***Formula***
Dissimilarity	$\sum_{i,j=0}^{M-1}(Pij\|i-j\|)$
Correlation	$\sum_{i,j=0}^{M-1} Pij\left[\frac{(i-\mu_i)(j-\mu_j)}{(\sigma_i^2)(\sigma_j^2)}\right]$
Homogeneity	$\sum_{i,j=0}^{M-1}\frac{Pij}{1+(i-j)^2}$
Contrast	$\sum_{i,j=0}^{M-1} Pij\,(i-j)^2$
ASM	$\sum_{i=0}^{M-1}\sum_{j=0}^{M-1} P(i,j)^2$
Energy	$\sqrt{ASM}$

The present study employed 4-Level of 2D-DWT decomposition and for each level of decomposition properties which their equations are given in Table 1 were extracted for each image (LH, HL, HH). Pixels distance was selected as 5 and for angles of [0, π/2, π/3, π/4, π/6, 3π/4], the statistical features were extracted. Thus, rather than accepting whole decomposition as a feature, more discriminative and numerical features were extracted. These extracted features are suitable for a ML algorithm since they represent enough statistical information about both cracked and non-cracked images.

However, the classical ML algorithms may not be sufficient for capturing the underlying function required to discriminate between cracked and non-cracked images. Moreover, since the primary objective of this study is to classify both the surface material and the presence of cracks, more advanced AI approaches may be better suited for the task at hand. For these reasons, an ANN structure was employed, as it offers a more flexible structure that can be modified to fit the needs of the problem. The following subsection provides an overview of the background knowledge necessary for understanding ANNs.

2.4. Artificial Neural Networks (ANNs)

An ANN is an AI method that can often effectively capture the underlying discriminative properties of a feature set and is commonly used for tasks such as classification. ANNs consist of multiple layers, each with a specific number of neurons. Each neuron has an activation function that is triggered when certain threshold value is reached. Each layer in the ANN is fully interconnected with the subsequent layer and is updated based on the computed error rate of the network's output. This process of error backpropagation enables the network to adjust its weights and biases in a manner that minimizes the overall loss or error of the model during training. Due to the aforementioned characteristics, ANNs have often been compared to the structure and function of the human brain, as they both involve the processing of complex information through the use of interconnected units. A simple ANN structure is given in **Figure 5**.

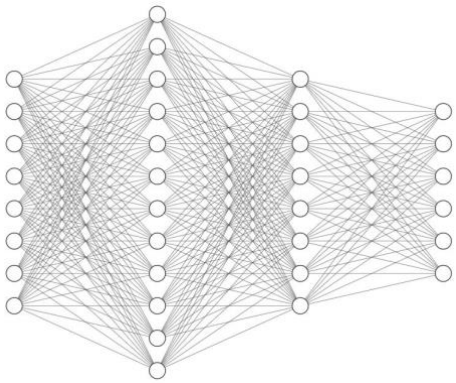


Figure 5. *An Example Artificial Neural Network Structure*

Figure 5 depicts the example architecture of an ANN with 8 input nodes and 6 output nodes, while the

intermediary layers are referred to as hidden layers and may have varying numbers of neurons. Each neuron has an activation function that ensures the non-linearity through the network. Because of their easily modifiable structure and performance in classification tasks, ANNs were chosen for this study.

Each of the mentioned methods has its own set of parameters that must be determined prior to conducting the experiment. Thus, the next section defines the experimental settings for this study. Also, it presents the general overview of the proposed method.

3. Experimental Settings and Overview of the Proposed Method

The dataset used in this study is generally employed for classifying images as either cracked or non-cracked. However, in addition to the binary classification of surface cracks, determining the type of surface is also crucial for further analysis and emergency responses. To address this, the class labels in this study were modified to include not only cracked and non-cracked labels, but also information regarding the type of surface. To make labels suitable for the ANN outputs, they were modified using the Label Encoding. The encoded labels are given in **Table 2**.

Table 2. *Labels Modification*

Label Name	*Original Label*	*Encoded Label*
Decks – Cracked	0	[1 0 0 0 0 0]
Decks – Non - Cracked	1	[0 1 0 0 0 0]
Pavements – Cracked	2	[0 0 1 0 0 0]
Pavements – Non - Cracked	3	[0 0 0 1 0 0]
Walls - Cracked	4	[0 0 0 0 1 0]
Walls – Non – Cracked	5	[0 0 0 0 0 1]

As mentioned earlier, the overall distribution of the dataset is imbalanced, with significantly more non-cracked images than images with cracks. This data imbalance could potentially lead to poor results when training a model to classify crack images. To address the data imbalance, 2500 images were selected for each label to balance the number of samples in the dataset. The feature extraction process involved a 2D-DWT with a 4-Level decomposition using the Bior1.5 wavelet function. For each level of decomposition, the GLCM with parameters as specified in Section 2.3, was utilized to extract statistical properties of the resulting Wavelet coefficients. Finally, an ANN structure was constructed, with 432 inputs, 6 hidden layers, and 6 outputs. The activation function used for each hidden layer was the Hyperbolic Tangent (Tanh), while the output layer employed the SoftMax activation function to generate the probability of each label. The BinaryCrossentropy loss function and Adam optimizer were decided as hyperparameters. To prevent overfitting of the ANN, dropout and regularization techniques were employed on various layers. All features were normalized to ensure faster convergence of the ANN. Finally, the dataset was divided into train and test sets, with a ratio of 80% for training and 20% for testing. Parameters of the methods are detailed in **Table 3**.

Table 3. *Parameters of the Methods*

Method / Settings	*Parameter(s)*
Number of Images for Each Label	2500
2D DWT	4 Level Decomposition. Bior1.5 Wavelet Function for each level.
GLCM	Distance of 5 Pixels. Angles of [0, $\pi/2$, $\pi/3$, $\pi/4$, $\pi/6$, $3\pi/4$]
ANN	6 Hidden Layers [512,256,128,256,64,32]
• Dropout	After 4^{th} and 6^{th} layers and ratio of 0.2 and 0.3 respectively.
• Regularization	First two layers have L1 and L2 regularization. The rest have L2 regularization except the final layer.
• Optimizer	Adam Optimizer – Learning Rate of 10^{-5}
• Number of Epoch	25
• Batch Size	8
• Loss Function	BinaryCrossentropy
• Train and Test set Ratio	80% - 20% respectively

3.1. Evaluation Metrics

The classical evaluation metrics for a classification problem were employed. The proposed method was evaluated in terms of accuracy, precision, recall, and F1 score. Additionally, the Area Under Curve (AUC) metric was employed, which serves as an indicator of the model's capacity to differentiate between classes. Moreover, the results of the method were presented in the form of a confusion matrix. The formulas for each metric are provided in **Table 4**.

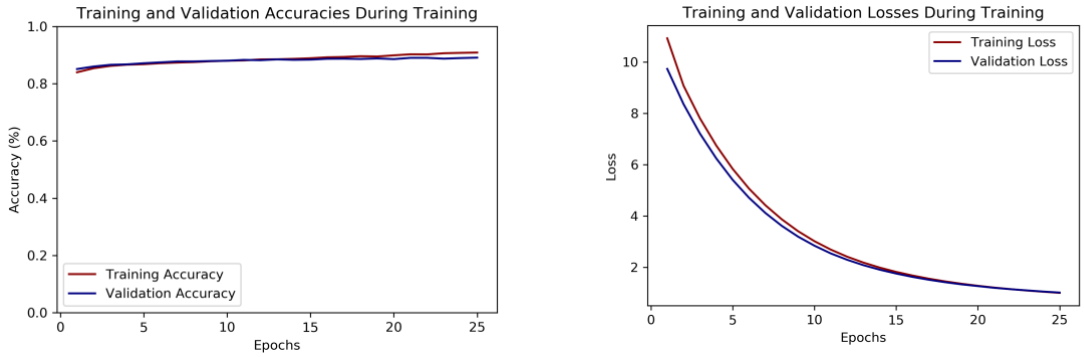
Table 4. *Evaluation Metrics*

Metric	***Formula***
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F_1 Score	$F_1 = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$

To clarify the abbreviations used in **Table 4**, we draw an analogy between surface types, specifically Wall, and two categories of inner classes: cracked and non-cracked. The abbreviation 'TP' stands for True Positive, which is the number of correct predictions made by the model. This means that the model predicts Wall-Crack and the sample is actually Wall-Crack. Similarly, 'TN' stands for True Negative, which is the number of negative predictions made by the model. In this case, the model predicts the sample as non-Wall and non-cracked, and the sample is actually non-Wall and non-cracked. On the other hand, 'FP' and 'FN' are abbreviations for False Positive and False Negative, respectively. For False Positive, the model predicts Wall-Crack when the sample is actually not Wall-Crack. Finally, False Negative represents the number of predictions where the model predicts the sample as non-Wall and non-cracked when the sample is actually Wall-Crack.

4. Experimental Results

We first inspect the results by examining the accuracy and loss metrics during training for both the train set and test set. The behavior of accuracy and loss metrics are given in Figure 6 side by side.



(a) (b)

Figure 6. *Accuracies and Losses Per Epoch. (a) Accuracy (b) Loss (Figure is in color in online version of paper)*

Figure 6 (a) shows that the accuracy of both the train set and the test set (used as the validation set) exhibited high performance. Although heavy regularizations were employed, the constructed ANN model exhibited signs of overfitting at the end of the training phase. However, the same deductions cannot be made for **Figure 6 (b)**, which shows the model's loss performance. Both the training and validation set losses tended to smoothly converge to a minimum. Final accuracy rates for both the train set and the validation set were given in **Table 5**.

Table 5. *Train and Test Set Performances of the Model*

Set	***Accuracy***
Train	92.19%
Test	89.12%

To gain a comprehensive understanding of the model, additional metrics as specified in **Table 4** were assessed on both the train and test sets following the completion of the model's training process. The overall results in terms of metrics are given **Table 6.**

Table 6. *The Overall Results for the Model*

Set	***Metric***	***Result***
Train	Precision	83.08%
	Recall	66.72%
	F1 Score	73.88%
	AUC	96.80%
Test	Precision	72.33%
	Recall	56.14%
	F1 Score	63.07%
	AUC	93.69%

Upon examining the results provided in **Table 6**, it is evident that the model exhibits relatively better performance on the train set as compared to the test set. One of the apparent reasons for this could be attributed to the number of samples used during the training phase of the model. Increasing the amount of data available could be one solution to this overfitting problem. It is known that providing an ANN structure with a larger amount of data often results in improved performance. The model's performance on precision was higher than other metrics which means that out of all the samples, 83% of them were classified correctly. The precision rate on the test set was 72.3%, which is a relatively good performance for such a challenging task. Finally, the model's performance was further evaluated using a confusion matrix, which is shown in **Figure 7**.

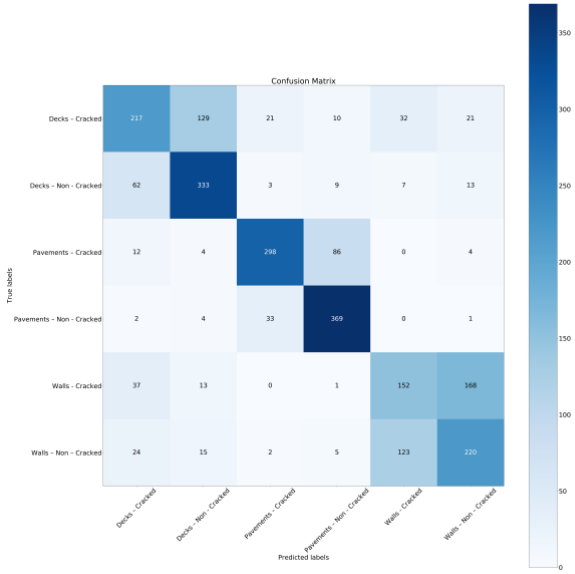


Figure 7. *Confusion Matrix for the Model (Figure is in color in online version of paper)*

Figure 7 highlights one of the primary factors that may have contributed to the relatively low performance of the model on the test set. Specifically, **Figure 7** indicates that while the model is generally effective at localizing surface materials, its ability to discriminate between inner-class categories appears to be not optimal. To provide a specific illustration, the model demonstrates a proficient capability in distinguishing between Wall and Pavement classes. However, the model appears to face considerable difficulty in effectively discerning between the cracked and non-cracked variants of the Wall surface category. As shown in **Figure 2**, hard examples of the dataset are one possible reason of dispersed inner class results in **Figure 7**. Despite the identified limitations, the findings suggest that the model's overall performance was satisfactory for the intended application.

To thoroughly assess the effectiveness of the proposed method, a quantitative analysis was conducted. This analysis involved a specific comparison of its performance against established methods documented in the existing literature, with a particular emphasis on CNNs. The comparison was based on the accuracy metric, as not all metrics presented in this paper were available in the literature within this context. Additionally, this comparison was limited to the SDNET 2018 dataset.

The methods introduced by Slonski [29] provide a comprehensive understanding of the behavior of Convolutional Neural Networks (CNNs) in the context of classification, specifically when applied to the SDNET 2018 dataset. Slonski's approach involved employing CNNs for the purpose of classifying cracks within the SDNET 2018. The method presented in this paper not only accomplishes crack classification but

also extends its capability to discern the specific type of surface. Consequently, our proposed method demonstrates superior performance compared to the conventional binary crack classification achieved in [29]. Another strategy put forth by Chianese et al. [30] involves Transfer Learning. Their approach was utilization of three pre-established CNN architectures: AlexNet, Inception-V3, and ResNet-101According to the methodology presented in [30], our approach increased the performance. A tabulated representation of this comparative analysis is provided in Table 7.

Table 7. *Comparison with CNNs Approaches*

Method	***Accuracy***
Slonski,M. [29] – From Scratch CNN	86.00%
Slonski,M. [29] – Pretrained VGG16 – Data Augmentation	88.00%
Chianese, R. et al [30] - AlexNet	87.30%
Chianese, R. et al [30] – Inception-V3	84.67%
Chianese, R. et al [30] – ResNet-101	86.00%
Proposed Method	89.12%

5. Conclusion

Structural deformation is a natural phenomenon that occurs in most structures over time. These deformations typically manifest in the exterior materials of structures, such as roads and buildings. Persistent monitoring of structural degradation is of paramount importance in ensuring the safety and reliability of structures. Failure to regularly inspect and address such deformations can pose a serious threat to both the economy and, more importantly, human lives. Furthermore, merely detecting cracks may not be sufficient for taking appropriate actions, as the type of surface material can also significantly influence the structural degradation process. Nevertheless, conducting regular inspections of structures demands a significant number of human resources, which is often impractical to implement in real-world scenarios.

Given this challenge, there is a need for more practical solutions to address the issue, including advancements in technology. In this regard, AI solutions are employed. The spectrum of AI methods, that can be categorized as ML algorithms, is vast and varies from classical ML algorithm to more advanced techniques such as CNNs. The classical ML algorithms work based on the features that are extracted from a given image. However, they may not produce acceptable performance since they are not flexible in modification. Also, they are bound by the quality of the extracted features. On the other hand, mechanisms that do not require human intervention, such as CNNs, are both computationally expensive and may produce ineffective features.

For these challenges in mind, the present study proposed an end-to-end ML approach for classifying the cracks and non-cracked images and material of the surface parallel. For the feature extraction part, 2D-DWT was utilized, as it is known that DWT enhances the frequency changes in an image. For the ML part, an ANN structure was employed since it provides a more flexible mechanism and allows us to control all aspects of it.

The study could be enhanced in several ways. One possible enhancement could be using different Wavelet functions to assess to performance of the model more. Another proper improvement is to be adding color feature to enhance the approach, since the color of each surface differs.

Different from the other studies in the literature, the current work classifies both cracked and non-cracked images and surface material at the same time. The results suggested that the proposed mechanism achieved an acceptable level of accuracy, although the task at hand is difficult.

Declaration of interest

The authors declares that there is no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the dataset provided by Utah State University.

Nomenclature

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Networks
CIP	Conventional Image Processing
CNN	Convolutional Neural Networks
DWT	Discrete Wavelet Transforms
GLCM	Gray Level Co-Occurence Matrix
NSGA	Hough Transform
HT	Hyperbolic Tangent

LBP	Local Binary Pattern
ML	Machine Learning
PCA	Principal Component Analysis
SVM	Support Vector Machines

References

[1] D. Ai, G. Jiang, S.-K. Lam, P. He, and C. Li, "Computer vision framework for crack detection of civil infrastructure—A review," *Engineering Applications of Artificial Intelligence*, 117 (2023) 10547; 10.1016/j.engappai.2022.105478.

[2] E. Mohammed Abdelkader, "On the hybridization of pre-trained deep learning and differential evolution algorithms for semantic crack detection and recognition in ensemble of infrastructures," *Smart and Sustainable Built Environment*, 11(3) (2022) 740–764; 10.1108/SASBE-01-2021-0010.

[3] L. Attard, C. J. Debono, G. Valentino, M. Di Castro, A. Masi, and L. Scibile, "Automatic crack detection using mask R-CNN," In: 11th international symposium on image and signal processing and analysis (ISPA), IEEE, (2019), 152–157.

[4] G. Lu, X. He, Q. Wang, F. Shao, J. Wang, and X. Zhao, "MSCNet: A Framework with a Texture Enhancement Mechanism and Feature Aggregation for Crack Detection," *IEEE Access*, 10 (2022) 26127–26139; 10.1109/ACCESS.2022.3156606.

[5] Z. Xu *et al.*, "Pavement crack detection from CCD images with a locally enhanced transformer network," *International Journal of Applied Earth Observation and Geoinformation*, 110 (2022) 102825; https://doi.org/10.1016/j.jag.2022.102825.

[6] P. Gupta and M. Dixit, "Image-based crack detection approaches: a comprehensive survey," *Multimedia Tools and Applications*, 81(28) (2022) 40181–40229; https://doi.org/10.1007/s11042-022-13152-z.

[7] L. Ali, F. Alnajjar, W. Khan, M. A. Serhani, and H. Al Jassmi, "Bibliometric analysis and review of deep learning-based crack detection literature published between 2010 and 2022," *Buildings*, 12(4) (2022) 432; https://doi.org/10.3390/buildings12040432.

[8] N. Safaei, O. Smadi, A. Masoud, and B. Safaei, "An automatic image processing algorithm based on crack pixel density for pavement crack detection and classification," *International Journal of Pavement Research and Technology*, 15(1) (2022) 159–172; https://doi.org/10.1007/s42947-021-00006-4.

[9] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, 57(2) (2018) 787–798; https://doi.org/10.1016/j.aej.2017.01.020.

[10] A. Ahmadi, S. Khalesi, and A. Golroo, "An integrated machine learning model for automatic road crack detection and classification in urban areas," *International Journal of Pavement Engineering*, 23(10) (2022) 3536–3552; https://doi.org/10.1080/10298436.2021.1905808.

[11] J. Illingworth and J. Kittler, "A survey of the hough transform," *Computer Vision, Graphics, and Image Processing*, 44(1) (1988) 87–116; https://doi.org/10.1016/S0734-189X(88)80033-1.

[12] C. D. Sutton, "Classification and regression trees, bagging, and boosting," *Handbook of statistics*, 24 (2005) 303–329; https://doi.org/10.1016/S0169-7161(04)24011-1.

[13] V. Jakkula, "Tutorial on support vector machine (svm)," *School of EECS, Washington State University*, 37(2.5) (2006).

[14] G. Zhang, X. Huang, S. Z. Li, Y. Wang, and X. Wu, "Boosting Local Binary Pattern (LBP)-Based Face Recognition," In: Chinese Conference on Biometric Recognition, (2014) 179–186.

[15] R. Bro and A. K. Smilde, "Principal component analysis," *Analytical Methods*, 6(9) (2014) 2812–2831; https://doi.org/10.1039/C3AY41907J.

[16] C. Chen, H. Seo, C. H. Jun, and Y. Zhao, "Pavement crack detection and classification based on fusion feature of LBP and PCA with SVM," *International Journal of Pavement Engineering*, 23(9) (2022) 3274–3283; https://doi.org/10.1080/10298436.2021.1888092.

[17] N.-D. Hoang, "Automatic detection of asphalt pavement raveling using image texture based feature extraction and stochastic gradient descent logistic regression" *Automation in Construction*, 105 (2019) 102843; https://doi.org/10.1016/j.autcon.2019.102843

[18] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3 (6) (1973) 610–621;10.1109/TSMC.1973.4309314.

[19] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, 3361 (10) (1995) 255-258, 1995.

[20] R. Rojas "The backpropagation algorithm," In: Neural Networks. Springer, Berlin, Heidelberg (1996) 149–182.

[21] C. Liu and B. Xu, "A night pavement crack detection method based on image-to-image translation," *Computer-Aided Civil and Infrastructure Engineering*, 37(13) (2022) 1737–1753; https://doi.org/10.1111/mice.12849.

[22] Q. Yang, W. Shi, J. Chen, and W. Lin, "Deep convolution neural network-based transfer learning method for civil infrastructure crack detection," *Automation in Construction*, 116 (2020) 103199; https://doi.org/10.1016/j.autcon.2020.103199.

[23] X. Zhang, D. Rajan, and B. Story, “Concrete crack detection using context-aware deep semantic segmentation network,” *Computer-Aided Civil and Infrastructure Engineering*, 34(11) (2019) 951–971; https://doi.org/10.1111/mice.12477.

[24] A. Chordia, S. Sarah, M. K. Gourisaria, R. Agrawal, and P. Adhikary, “Surface Crack Detection Using Data Mining and Feature Engineering Techniques,” In: IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), (2021) 1–7.

[25] L. Cong, J. Shi, T. Wang, F. Yang, and T. Zhu, “A method to evaluate the segregation of compacted asphalt pavement by processing the images of paved asphalt mixture,” *Construction and Building Materials*, 224 (2019) 622–629; https://doi.org/10.1016/j.conbuildmat.2019.07.041.

[26] S. Dorafshan, R. J. Thomas, and M. Maguire, “SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks,” *Data in brief*, 21 (2018) 1664–1668; https://doi.org/10.1016/j.dib.2018.11.015.

[27] M. Mirbod and M. Shoar, “Intelligent Concrete Surface Cracks Detection using Computer Vision, Pattern Recognition, and Artificial Neural Networks,” *Procedia Computer Science*, 217 (2023) 52–61; https://doi.org/10.1016/j.procs.2022.12.201.

[28] I. Daubechies, “Ten lectures on wavelets”, SIAM, 1992.

[29] M. Sloński, “A comparison of deep convolutional neural networks for image-based detection of concrete surface cracks,” *Computer assisted methods in Engineering and Science*, 26 (2) (2019) 105–112; http://dx.doi.org/10.24423/cames.267.

[30] R. Chianese, A. Nguyen, V. Gharehbaghi, T. Aravinthan, and M. Noori, “Influence of image noise on crack detection performance of deep convolutional neural networks,” In: Proceedings of the 10th International Conference on Structural Health Monitoring of Intelligent Infrastructure (SHMII 10), (2021) 1681 – 1688.